View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Elsevier - Publisher Connector

ELSEVIER Linear Algebra and its Applications 335 (2001) 189–204

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

DGMRES: A GMRES-type algorithm for Drazin-inverse solution of singular nonsymmetric linear systems

Avram Sidi

Computer Science Department, Technion, Israel Institute of Technology, Haifa 32000, Israel

Received 20 January 2000; accepted 14 February 2001

Submitted by M. Neumann

Abstract

In a recent work by the author [Linear Algebra Appl. 298 (1999) 99] Krylov subspace methods were derived for the Drazin-inverse solution of consistent or inconsistent linear systems of the form $Ax = b$, where $A \in \mathbb{C}^{N\times N}$ is a singular and in general non-hermitian matrix that has an arbitrary index. One of these methods, modeled after the generalized conjugate residual method (GCR) and denoted DGCR, is considered in the present work again. It is shown that all of the approximations produced by DGCR exist, and a GMRES-like algorithm, denoted DGMRES, for its implementation is derived. Like GMRES, DGMRES too is economical computationally and storagewise. © 2001 Elsevier Science Inc. All rights reserved.

AMS classification: 15A06; 15A09; 65F10; 65F50

Keywords: Singular linear systems; Drazin-inverse solution; Projection methods; Krylov subspace methods

1. Introduction

Consider the linear system

$$Ax = b, \tag{1.1}$$

where $A \in \mathbb{C}^{N\times N}$ is a singular matrix and ind(A) is arbitrary. Here ind(A), the index of A, is the size of the largest Jordan block corresponding to the zero eigenvalue of A. In a recent paper, [13], we presented a general framework within which Krylov subspace methods for the Drazin-inverse solution of (1.1) can be developed and their

E-mail address: asidi@cs.technion.ac.il (A. Sidi).

0024-3795/01/$ - see front matter © 2001 Elsevier Science Inc. All rights reserved.
PII: S 0 0 2 4 - 3 7 9 5 (0 1) 0 0 2 8 9 - 0

properties discussed. Here we recall that the Drazin-inverse solution of (1.1) is the vector $A^{\mathrm{D}}b$, where A^{D} denotes the Drazin inverse of the singular matrix A. For the Drazin inverse and its properties, see, e.g., [2] or [4].

In [13] we do not put any restrictions on the matrix A. Thus, in general, A is non-hermitian, $\mathrm{ind}(A)$ is arbitrary, and the spectrum of A can have any shape. Neither do we put any restrictions on the linear system (1.1). This system may be consistent or inconsistent. We only assume that $\mathrm{ind}(A)$ is known. (By going through the details of [13] it becomes clear that the methods developed there remain valid also when $\mathrm{ind}(A)$ is overestimated.)

It seems that Ref. [13] is the first paper that treats the singular system (1.1) under the most general conditions mentioned in the previous paragraph. In contrast, other papers in the literature treat either the cases in which (1.1) is consistent or it is inconsistent but A is hermitian, and in all cases $\mathrm{ind}(A) = 1$. For a review of the relevant literature, we refer the reader to [13]. Here we will only note the work of Calvetti et al. [3] that deals with inconsistent singular systems (1.1) with A hermitian, in which two efficient conjugate gradient type algorithms for such problems are presented. We shall also mention the work of Sidi and Kluzner [14] that gives a Lanczos [11] type method and a bi-conjugate gradient (Bi-CG)-type algorithm for it in the general case that we treat in the present work. For Bi-CG see [8]. Finally, we mention that the case of the arbitrary index has also been treated by Climent et al. [5] within the context of semi-iterative methods in the presence of real non-negative spectra, which generalizes the work of Hanke and Hochbruck [10] that concerns $\mathrm{ind}(A) = 1$.

In [13] we derived several Krylov subspace methods for computing $A^{\mathrm{D}}b$, one of them being analogous to the generalized conjugate residual method (GCR) of Eisenstat et al. [7] and we denoted it DGCR for short. Now the best implementation of GCR is via GMRES of Saad and Schulz [12] that has been used successfully in many applications. The purpose of the present work is to present an implementation of DGCR for singular systems that is analogous to GMRES for non-singular systems. This GMRES-like implementation, just like GMRES itself, is stable numerically and economical computationally and storagewise. We shall denote this implementation DGMRES for short.

The plan of this paper is as follows. In Section 2 we shall give a brief review of the major developments of [13] relevant to us here. We shall discuss the projection method approach to (1.1) in general, and DGCR in particular. In Section 3 we shall recall the finite termination property and prove in addition that DGCR terminates successfully unconditionally. We shall also recall the upper bounds on the errors in the approximations produced by DGCR. In addition, we shall prove that the approximations to $A^{\mathrm{D}}b$ produced by DGCR exist and are unique unconditionally. The development of DGMRES, the implementation of DGCR, is the subject of Section 4. We design DGMRES such that when we set $\mathrm{ind}(A) = 0$ throughout, DGMRES reduces to GMRES. In this sense DGMRES is an extension of GMRES that achieves the Drazin-inverse solution of singular systems. In Section 5 we summarize the steps of DGMRES and in Section 6 we propose an effective mode of usage for DGMRES

that we denote DGMRES(m) and that is analogous to GMRES(m) and that requires a fixed amount of storage for its implementation. Finally, in Section 7 we demonstrate the use of DGMRES with numerical examples.

Before we end this section we would like to note that the approach of the work [13] to Krylov subspace methods can be extended to derive methods analogous to Bi-CGSTAB (see [15]) and QMR (see [9]) for the singular systems considered in this work. We plan to come back to this subject in a future publication.

2. Review of projection methods for $A^{\mathrm{D}}b$

The methods we are interested in start with an arbitrary initial vector x_0 and generate a sequence of vectors $x_1, x_2, \ldots$, according to

$$x_m = x_0 + q_{m-1}(A)r_0, \quad r_0 = b - Ax_0, \tag{2.1}$$

where $q_{m-1}(\lambda)$ is a polynomial in λ of degree at most $m-1$, given by

$$q_{m-1}(\lambda) = \sum_{i=1}^{m-a} c_i \lambda^{a+i-1}, \quad a = \mathrm{ind}(A). \tag{2.2}$$

Let us define

$$p_m(\lambda) = 1 - \lambda q_{m-1}(\lambda) = 1 - \sum_{i=1}^{m-a} c_i \lambda^{a+i}. \tag{2.3}$$

We call $p_m(\lambda)$ the mth residual polynomial since

$$r_m = b - Ax_m = [I - Aq_{m-1}(A)]r_0 = p_m(A)r_0. \tag{2.4}$$

Note that

$$p_m(0) = 1 \quad \text{and} \quad p_m^{(i)}(0) = 0, \quad i = 1, 2, \ldots, a. \tag{2.5}$$

The conditions in (2.5) are due to Eiermann et al. [6].

For convenience we denote by Π_m the class of polynomials of degree at most m and define

$$\Pi_m^0 = \{p \in \Pi_m : p(0) = 1 \text{ and } p^{(i)}(0) = 0,\ i = 1, \ldots, a\}. \tag{2.6}$$

Thus, the polynomial p_m described above is in Π_m^0.

The projection methods of [13] are now defined by demanding that the vector $A^a r_m$ be orthogonal to a given subspace W of dimension $m-a$. If we denote also by W the $N \times (m-a)$ matrix whose columns span the subspace W, then this orthogonality demand is equivalent to $W^* A^a r_m = 0$. As we have $r_m = r_0 - \sum_{i=1}^{m-a} c_i A^{a+i} r_0$ from (2.4), $W^* A^a r_m = 0$ amounts to the requirement that $c_1, \ldots, c_{m-a}$ satisfy the linear system

$$W^* A^{a+1} V c = W^* A^a r_0, \tag{2.7}$$

where $V \in \mathbb{C}^{N\times(m-a)}$ and $c \in \mathbb{C}^{m-a}$ are given by

$$V = [A^a r_0 | A^{a+1} r_0 | \cdots | A^{m-1} r_0] \quad \text{and} \quad c = [c_1, \ldots, c_{m-a}]^{\mathrm{T}}. \tag{2.8}$$

We see that a unique solution for c exists provided $\det(W^* A^{a+1} V) \neq 0$, and when it does we have $x_m = x_0 + V(W^* A^{a+1} V)^{-1} W^* A^a r_0$.

Different justifications for the validity of the projection approach are given in [13]. We mention one of them below. Before we can do that we need a few definitions.

We shall denote by $\hat{\mathcal{S}}$ the direct sum of the invariant subspaces of A corresponding to its non-zero eigenvalues, and by $\tilde{\mathcal{S}}$, its invariant subspace corresponding to its zero eigenvalue. Thus $\hat{\mathcal{S}}$ is $\mathcal{R}(A^a)$, the range of A^a, and $\tilde{\mathcal{S}}$ is $\mathcal{N}(A^a)$, the null space of A^a. Every vector in $\mathbb{C}^N$ can be expressed as the sum of two unique vectors, one in $\hat{\mathcal{S}}$ and the other in $\tilde{\mathcal{S}}$.

Definition 2.1. Let A be singular and $\mathrm{ind}(A) = a$, and let $\hat{u} \in \hat{\mathcal{S}}$ be given. Then a polynomial $P(\lambda)$ will be called the minimal a-incomplete polynomial of A with respect to the vector $\hat{u}$ if $P \in \Pi_m^0$ and m is smallest possible such that $P(A)\hat{u} = 0$.

The following result is Theorem 4.1 in [13] and will be used later in this work.

Theorem 2.1. *$P(\lambda)$ exists and is unique. Furthermore, its degree m satisfies $q \leqslant m \leqslant q + a$, where q is the degree of the minimal polynomial of A with respect to $\hat{u}$.*

The following result that is the justification of the projection approach alluded to above is Theorem 4.2 in [13].

Theorem 2.2. *Let $x_0 = \hat{x}_0 + \tilde{x}_0$, where $\hat{x} \in \hat{\mathcal{S}}$ and $\tilde{x}_0 \in \tilde{\mathcal{S}}$, be the initial vector in the projection method for computing $A^{\mathrm{D}}b$. Let also $P(\lambda)$ be the minimal a-incomplete polynomial of A with respect to $\hat{x}_0 - A^{\mathrm{D}}b$, and let m be its degree. Finally, let x_m be the vector generated by the projection method through* (2.1)–(2.8), *with* $\det(W^* A^{a+1} V) \neq 0$. *Then $x_m = A^{\mathrm{D}}b + \tilde{x}_0$.*

From Theorem 2.2 it is clear that the projection methods above terminate successfully in a finite number of steps, this number being N at most. If we pick $x_0 = 0$, which forces $\tilde{x}_0 = 0$ as well, they produce the Drazin-inverse solution $A^{\mathrm{D}}b$ to (1.1) upon termination.

3. Properties of DGCR

The projection method that we have called DGCR is defined by picking $W = A^{a+1}V$. The vector x_m produced by DGCR then can be shown to satisfy

$$\|A^a r_m\| = \min_{x \in x_0 + \mathcal{K}_{m-a}(A; A^a r_0)} \|A^a (b - Ax)\|, \tag{3.1}$$

where $\mathscr{K}_s(A; b) = \text{span}\{b, Ab, \ldots, A^{s-1}b\}$ is a Krylov subspace. The normal equations for the unknown coefficients $c_1, \ldots, c_{m-a}$ in (2.1) and (2.2) turn out to be precisely the projection equations given in (2.7) and (2.8).

Theorem 3.1. *The vector x_m exists uniquely and unconditionally for all $m \leqslant m_0$, m_0 being the degree of the minimal a-incomplete polynomial of A with respect to $\hat{x}_0 - A^{\mathrm{D}}b \in \hat{\mathscr{S}}$. Furthermore, $x_m = \hat{x}_m + \tilde{x}_m$ with $\hat{x}_m \in \hat{\mathscr{S}}$ and $\tilde{x}_m = \tilde{x}_0$ for all m.*

Proof. To prove the first part we have to show that $\det(W^*A^{a+1}V) \neq 0$ is satisfied automatically when $m \leqslant m_0$. Now by the fact that $W = A^{a+1}V$ it follows that $W^*A^{a+1}V = (A^{a+1}V)^*(A^{a+1}V)$ is non-singular provided $A^{a+1}V$ has full rank. By (2.8), $A^{a+1}V$ will have full rank provided its columns $A^{2a+1}r_0$, $A^{2a+2}r_0, \ldots, A^{a+m}r_0$ are linearly independent or, equivalently, provided the linear system

$$\sum_{i=1}^{m-a} d_i A^{2a+i} r_0 = 0 \tag{3.2}$$

has only the trivial solution $d_1 = \cdots = d_{m-a} = 0$. Now $r_0 = \hat{r}_0 + \tilde{r}_0$, where $\hat{r}_0 = \hat{b} - A\hat{x}_0 = -A(\hat{x}_0 - A^{\mathrm{D}}b) \in \hat{\mathscr{S}}$ and $\tilde{r}_0 = \tilde{b} - A\tilde{x}_0 \in \tilde{\mathscr{S}}$. Here we have used the fact that $AA^{\mathrm{D}}b = \hat{b}$. By the fact that $A^a v = 0$ for $v \in \tilde{\mathscr{S}}$, the equations in (3.2) become

$$A^{2a+2} \sum_{i=1}^{m-a} d_i A^{i-1}(\hat{x}_0 - A^{\mathrm{D}}b) = 0. \tag{3.3}$$

Since the summation in (3.3) is in $\hat{\mathscr{S}}$, we also have

$$\sum_{i=1}^{m-a} d_i A^{i-1}(\hat{x}_0 - A^{\mathrm{D}}b) = 0. \tag{3.4}$$

That is to say, the polynomial $\sum_{i=1}^{m-a} d_i \lambda^{i-1} \in \Pi_{m-a-1}$ is divisible by the minimal polynomial of A with respect to $\hat{x}_0 - A^{\mathrm{D}}b$, whose degree is q. Thus, $m - a - 1 \geqslant q$, hence $m \geqslant q + a + 1$, which contradicts the fact that $m \leqslant m_0 \leqslant q + a$ that follows from Theorem 2.1. This implies that the only solution to (3.4) is $d_1 = \cdots = d_{m-a} = 0$. This completes the proof of the first part.

To prove the second part we start by observing that, due to the fact that the polynomial $q_{m-1}(\lambda)$ in (2.1) is of the form $q_{m-1}(\lambda) = \lambda^a \sum_{i=1}^{m-a} c_i \lambda^{i-1}$, there holds $q_{m-1}(A)\tilde{r}_0 = 0$. Therefore, x_m in (2.1) can actually be written in the form $x_m = \hat{x}_m + \tilde{x}_m$ with $\hat{x}_m = \hat{x}_0 + q_{m-1}(A)\hat{r}_0 \in \hat{\mathscr{S}}$ while $\tilde{x}_m = \tilde{x}_0 \in \tilde{\mathscr{S}}$ for all m. This completes the proof. □

As has been shown in [13], $\|A^a r_m\|$ is a true norm for the error $\hat{x}_m - A^{\mathrm{D}}b$, where we know that $x_m = \hat{x}_m + \tilde{x}_0$, with $\hat{x}_m \in \hat{\mathscr{S}}$. From (3.1) it follows that

$$\|A^a r_a\| > \|A^a r_{a+1}\| > \|A^a r_{a+2}\| > \cdots > \|A^a r_{m_0}\| = 0, \tag{3.5}$$

where m_0 is the degree of the minimal a-incomplete polynomial of A with respect to $\hat{x}_0 - A^{\mathrm{D}}b$. Theorem 6.1 in [13], in addition, provides an upper bound on $\|A^a r_m\|$. We give the exact statement of this theorem here for completeness.

Theorem 3.2. *Denote the spectrum of A by $\sigma(A)$ and choose Ω to be a closed domain in the λ-plane that contains $\sigma(A)\backslash\{0\}$ but not $\lambda = 0$, such that its boundary is twice differentiable with respect to arclength. Denote by $\Phi(\lambda)$ the conformal mapping of the exterior of Ω onto the exterior of the unit disk $\{w : |w| \geqslant 1\}$. Then the vector x_m generated by DGCR satisfies*

$$\|A^a r_m\| \leqslant K\, m^{a+2(\hat{k}-1)}\rho^m \quad \textit{for all } m, \tag{3.6}$$

where K is a positive constant independent of m, $\hat{k} = \max\{k_j : k_j = \mathrm{ind}(A - \lambda_j I)$, $\lambda_j \in \sigma(A)\backslash\{0\}\}$, and $\rho = 1/|\Phi(0)| < 1$.

Clearly, the upper bound on $\|A^a r_m\|$ given in this theorem decreases to zero essentially exponentially in m due to the presence of the factor ρ^m. Needless to say, the result in (3.6) would be sharper when Ω is picked to be as small as possible as this reduces the value of ρ.

4. Implementation of DGCR by DGMRES

The most straightforward way of determining x_m would be by computing $x_m = x_0 + \sum_{i=1}^{m-a} c_i A^{a+i-1} r_0$, with the c_i determined by solving the linear equations in (2.7). For DGCR these equations are nothing but the normal equations that result from (3.1), as we mentioned before. Even though this process is economical computationally and requires the storage of only the $m + 2$ vectors x_0 and $A^{a+i}r_0$, $i = 0, 1, \ldots, m$, it is well known that solving least squares problems via the normal equations is not the best procedure to follow. In addition, the fact that the power iterations $A^k r_0$ become less and less linearly independent may cause the matrix $A^{a+1}V = [A^{2a+1}r_0|A^{2a+2}r_0|\cdots|A^{m+a}r_0]$ to be more and more ill conditioned, causing the numerical solution for the c_i to be not very accurate.

For these reasons we now turn to an altogether different implementation of DGCR. The algorithm that we develop here is analogous to GMRES, the most efficient implementation of GCR for non-singular systems. We must note, however, that, despite the analogy, DGMRES turns out to be quite different from GMRES due to the inherent mathematical differences between DGCR and GCR.

As $x_m = x_0 + \sum_{i=1}^{m-a} c_i A^{a+i-1} r_0$, we start by orthogonalizing the Krylov vectors $A^a r_0, A^{a+1} r_0, \ldots$, by the Arnoldi–Gram–Schmidt process, see [1,12], carried out numerically like the modified Gram–Schmidt process:

1. Let $\beta = \|A^a r_0\|$ and set $v_1 = \beta^{-1}(A^a r_0)$.

2. For $i = 1, 2, \ldots,$ do
   Compute $h_{ji} = (v_j, Av_i),\ j = 1, \ldots, i.$
   Compute $\hat{v}_i = Av_i - \sum_{j=1}^{i} v_j h_{ji}.$
   Let $h_{i+1,i} = \|\hat{v}_i\|$ and set $v_{i+1} = \hat{v}_i / h_{i+1,i}.$

Consequently, we have a set of orthonormal vectors $v_1, v_2, \ldots,$ that satisfies

$$Av_i = \sum_{j=1}^{i+1} v_j h_{ji}, \quad i = 1, 2, \ldots, \tag{4.1}$$

as long as $i \leqslant q - 1$, where q is the degree of the minimal polynomial of A with respect to $A^a r_0$, hence with respect to v_1. This is the same as requiring that $Av_i \notin$ span $\{v_1, \ldots, v_i\}$. Furthermore, for each k,

$$\begin{aligned} \text{span}\{v_1, v_2, \ldots, v_k\} &= \text{span}\{A^a r_0, A^{a+1} r_0, \ldots, A^{k+a-1} r_0\} \\ &= \mathscr{K}_k(A; A^a r_0). \end{aligned} \tag{4.2}$$

If we now define the $N \times k$ matrix $\hat{V}_k$ by

$$\hat{V}_k = [v_1 | v_2 | \cdots | v_k], \quad k = 1, 2, \ldots, \tag{4.3}$$

then we can write for $m \leqslant m_0$

$$x_m = x_0 + \hat{V}_{m-a} \xi_m \quad \text{for some } \xi_m \in \mathbb{C}^{m-a}, \tag{4.4}$$

and we need to determine ξ_m. First, by $r_m = r_0 - A\hat{V}_{m-a}\xi_m$, we have

$$A^a r_m = A^a r_0 - A^{a+1}\hat{V}_{m-a}\xi_m = \beta v_1 - A^{a+1}\hat{V}_{m-a}\xi_m. \tag{4.5}$$

Next, provided $k \leqslant q - 1$, from (4.1) we can write

$$A\hat{V}_k = \hat{V}_{k+1}\bar{H}_k; \quad \bar{H}_k = \begin{bmatrix} h_{11} & h_{12} & \ldots & \ldots & h_{1k} \\ h_{21} & h_{22} & \ldots & \ldots & h_{2k} \\ 0 & h_{32} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & h_{kk} \\ 0 & \ldots & \ldots & 0 & h_{k+1,k} \end{bmatrix}. \tag{4.6}$$

Note that $\bar{H}_k \in \mathbb{C}^{(k+1)\times k}$. Now $\text{rank}(A\hat{V}_k) = k$ since the vectors $Av_1, Av_2, \ldots, Av_k$ are linearly independent when $k \leqslant q - 1$. Since also $\text{rank}(\hat{V}_{k+1}) = k + 1$, and $\text{rank}(A\hat{V}_k) \leqslant \min\{\text{rank}(\hat{V}_{k+1}), \text{rank}(\bar{H}_k)\}$, we have that $\text{rank}(\bar{H}_k) = k$. That is, $\bar{H}_k$ has full rank. Next, if we apply (4.6) to $A^{a+1}\hat{V}_{m-a}$, provided $m \leqslant q - 1$, we have

$$\begin{aligned} A^{a+1}\hat{V}_{m-a} &= A^a \hat{V}_{m-a+1}\bar{H}_{m-a} \\ &= A^{a-1}\hat{V}_{m-a+2}\bar{H}_{m-a+1}\bar{H}_{m-a} = \cdots = \hat{V}_{m+1}\hat{H}_m; \\ \hat{H}_m &\equiv \bar{H}_m \bar{H}_{m-1} \cdots \bar{H}_{m-a}. \end{aligned} \tag{4.7}$$

Consequently, provided $m \leqslant q - 1$,

$$A^a r_m = \beta v_1 - \hat{V}_{m+1}\hat{H}_m\xi_m, \tag{4.8}$$

and since $v_1 = \hat{V}_{m+1}e_1$, where $e_1 = [1, 0, 0, \ldots, 0]^{\mathrm{T}} \in \mathbb{C}^{m+1}$, and $\hat{V}^*_{m+1}\hat{V}_{m+1} = I_{(m+1)\times(m+1)}$, we finally have

$$\|A^a r_m\| = \|\hat{V}_{m+1}(\beta e_1 - \hat{H}_m\xi_m)\| = \|\beta e_1 - \hat{H}_m\xi_m\|, \tag{4.9}$$

the l_2-norm on the right-hand side of (4.9) now being defined in $\mathbb{C}^{m+1}$. Note that $\hat{H}_m \in \mathbb{C}^{(m+1)\times(m-a)}$ and it has full rank. To see this we observe that rank $(A^{a+1}\hat{V}_{m-a}) = m - a$ and that $\mathrm{rank}(A^{a+1}\hat{V}_{m-a}) \leqslant \min\{\mathrm{rank}(\hat{V}_{m+1}), \mathrm{rank}(\hat{H}_m)\}$, so that $\mathrm{rank}(\hat{H}_m) = m - a$.

We have thus reduced the $N \times (m - a)$ least squares problem of (3.1) to the $(m + 1) \times (m - a)$ least squares problem

$$\|A^a r_m\| = \|\beta e_1 - \hat{H}_m\xi_m\| = \min_{\xi\in\mathbb{C}^{m-a}} \|\beta e_1 - \hat{H}_m\xi\|. \tag{4.10}$$

Note that N is normally very large and $m \ll N$, which implies that the problem in (4.10) is very small. We can simplify the solution for ξ_m further by applying the QR factorization to $\hat{H}_m$. Thus, $\hat{H}_m = Q_m R_m$, where $Q_m \in \mathbb{C}^{(m+1)\times(m-a)}$ is a unitary matrix in the sense that $Q^*_m Q_m = I_{(m-a)\times(m-a)}$ and $R_m \in \mathbb{C}^{(m-a)\times(m-a)}$ is upper triangular. Since $\hat{H}_m$ has full rank, R_m is also non-singular. Thus, ξ_m in (4.10) can be obtained by solving the upper triangular system

$$R_m\xi_m = \beta(Q^*_m e_1). \tag{4.11}$$

Since $\bar{H}_k$ is a $(k + 1) \times k$ upper Hessenberg-type matrix by (4.6), $\hat{H}_m$ is an $(m + 1) \times (m - a)$ matrix of similar form:

$$\hat{H}_m = \begin{bmatrix}
\hat{h}_{11} & \hat{h}_{12} & \cdots & \cdots & \cdots & \cdots & \hat{h}_{1,m-a} \\
\hat{h}_{21} & \hat{h}_{22} & \cdots & \cdots & \cdots & \cdots & \hat{h}_{2,m-a} \\
\vdots & \vdots & \ddots & \ddots & & & \vdots \\
\hat{h}_{a+2,1} & \vdots & & \ddots & \ddots & & \vdots \\
0 & \hat{h}_{a+3,2} & & & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & & & \ddots & \hat{h}_{m-a-1,m-a} \\
\vdots & & \ddots & \ddots & & & \hat{h}_{m-a,m-a} \\
\vdots & & & \ddots & \ddots & & \vdots \\
\vdots & & & & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & 0 & \hat{h}_{m+1,m-a}
\end{bmatrix}. \tag{4.12}$$

Consequently, Q_m is also an $(m + 1) \times (m - a)$ matrix and has the same zero structure as $\hat{H}_m$.

Furthermore, $\hat{H}_{m+1}$ can be obtained as a simple update of $\hat{H}_m$ by first appending a row of zeros at the bottom of $\hat{H}_m$ and following that by appending the $(m + 2)$-vector $[\hat{h}_{1,m+1-a}, \ldots, \hat{h}_{m+2,m+1-a}]^{\mathrm{T}}$ as the $(m + 1 - a)$th column. To prove this

assertion and to show how the last column of $\hat{H}_{m+1}$ is determined we proceed as follows.

Let us define

$$\begin{aligned} G_k^{(0)} &= \bar{H}_k, \quad k = 1, 2, \ldots, \\ G_k^{(j)} &= \bar{H}_k G_{k-1}^{(j-1)}, \quad k = j+1, j+2, \ldots, \quad j = 1, 2, \ldots \end{aligned} \tag{4.13}$$

Of course, $G_k^{(j)} \in \mathbb{C}^{(k+1)\times(k-j)}$. Also, $G_m^{(a)} = \hat{H}_m$ for each $m \geqslant a+1$. What we need to show now is that

$$G_{m+1}^{(j)} = \left[\begin{array}{c|c} G_m^{(j)} & g_{m+1}^{(j)} \\ \hline 0_{m-j}^{\mathrm{T}} & \alpha_{m+1}^{(j)} \end{array}\right], \quad m \geqslant j+1, \tag{4.14}$$

where, of course, $g_{m+1}^{(j)} \in \mathbb{C}^{m+1}$, 0_k denotes the k-dimensional (column) zero vector, and $\alpha_{m+1}^{(j)}$ is a scalar. We do this by induction on j. The assertion is true for $j = 0$ by (4.13) and (4.6). Let us prove that it is true for $j = s+1$, assuming that it is true for $j = 1, \ldots, s$. By this assumption and by the definition of the $G_k^{(j)}$ in (4.13) we have

$$\begin{aligned} G_{m+1}^{(s+1)} &= \bar{H}_{m+1} G_m^{(s)} \\ &= \left[\begin{array}{c|c} \bar{H}_m & g_{m+1}^{(0)} \\ \hline 0_m^{\mathrm{T}} & \alpha_{m+1}^{(0)} \end{array}\right] \left[\begin{array}{c|c} G_{m-1}^{(s)} & g_m^{(s)} \\ \hline 0_{m-s-1}^{\mathrm{T}} & \alpha_m^{(s)} \end{array}\right], \quad m \geqslant s+2. \end{aligned} \tag{4.15}$$

From this we see that

$$G_{m+1}^{(s+1)} = \left[\begin{array}{c|c} G_m^{(s+1)} & g_{m+1}^{(s+1)} \\ \hline 0_{m-s-1}^{\mathrm{T}} & \alpha_{m+1}^{(s+1)} \end{array}\right], \quad m \geqslant s+2 \tag{4.16}$$

with

$$g_{m+1}^{(s+1)} = \bar{H}_m g_m^{(s)} + \alpha_m^{(s)} g_{m+1}^{(0)} \quad \text{and} \quad \alpha_{m+1}^{(s+1)} = \alpha_{m+1}^{(0)} \alpha_m^{(s)}. \tag{4.17}$$

We have thus proved the assertion. We have also shown via (4.17) how the last column of the matrix $G_{m+1}^{(s+1)}$ is obtained from $G_m^{(s)}$ in a simple way.

The fact that $\hat{H}_{m+1}$ is obtained from $\hat{H}_m$ by appending to $\hat{H}_m$ a row of zeros and an $(m+2)$-dimensional column vector implies that the matrices Q_{m+1} and R_{m+1} in the QR factorization of $\hat{H}_{m+1}$ can be obtained by updating Q_m and R_m in exactly the same way. This enables us to obtain the norms $\|A^a r_m\|$, $m \geqslant a$, without actually having to form x_m and r_m. Since from (4.9) and (4.11) there holds

$$\|A^a r_m\| = \|\beta e_1 - Q_m R_m \xi_m\| = \|\beta e_1 - \beta Q_m Q_m^* e_1\|, \tag{4.18}$$

we obtain after some manipulation that

$$\|A^a r_m\| = \beta\sqrt{1 - \|Q_m^* e_1\|^2}. \tag{4.19}$$

We have so far considered the determination of x_m with $m \leqslant q-1$ for which the algorithm above will never fail. Since in practice q will be very close to $N-a$ and hence to N, and m is to be picked much smaller than N, this algorithm will be sufficient for all practical purposes. Nevertheless, it is interesting to look at the cases in which $m \geqslant q$. For these cases (4.7) undergoes some changes that we need to study.

Let us take $m = q$ first. In this case (4.1) becomes

$$Av_q = \sum_{j=1}^{q} v_j h_{jq}, \tag{4.20}$$

as a result of which (4.6) becomes

$$A\hat{V}_q = \hat{V}_q H_q; \quad H_k = \begin{bmatrix} h_{11} & h_{12} & \dots & \dots & h_{1k} \\ h_{21} & h_{22} & \dots & \dots & h_{2k} \\ 0 & h_{32} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & h_{k,k-1} & h_{kk} \end{bmatrix}. \tag{4.21}$$

Note that $H_k \in \mathbb{C}^{k\times k}$ and has full rank. This changes (4.7) to

$$A^{a+1}\hat{V}_{q-a} = \hat{V}_q H_q \bar{H}_{q-1}\bar{H}_{q-2}\cdots\bar{H}_{q-a} \equiv \hat{V}_q \hat{H}_q, \tag{4.22}$$

with $\hat{H}_q \in \mathbb{C}^{q\times(q-a)}$. In addition, $\hat{H}_q$ has full rank.

For $m = q+1$ we similarly obtain

$$A^{a+1}\hat{V}_{q+1-a} = \hat{V}_q H_q^2 \bar{H}_{q-1}\bar{H}_{q-2}\cdots\bar{H}_{q+1-a} \equiv \hat{V}_q \hat{H}_{q+1}, \tag{4.23}$$

with $\hat{H}_{q+1} \in \mathbb{C}^{q\times(q+1-a)}$. $\hat{H}_{q+1}$ has full rank too. The way we continue is now clear. We see that the algorithm will not fail as the matrices R_m in the QR factorization of $\hat{H}_m$, $m \geqslant q$, will all be non-singular.

Finally, for $m = q+a$ we will have

$$A^{a+1}\hat{V}_q = \hat{V}_q H_q^{a+1} \equiv \hat{V}_q \hat{H}_{q+a}, \tag{4.24}$$

with $\hat{H}_{q+a} \in \mathbb{C}^{q\times q}$. Also, $\hat{H}_{q+a}$ has full rank. This means that the solution that is returned by $R_{q+a}\xi_{q+a} = \beta(Q_{q+a}^* e_1)$ actually gives $\|A^a r_{q+a}\| = 0$ as desired.

5. Summary of DGMRES

For the sake of clarity we now summarize the steps of DGMRES for the solution of a linear system $Ax = b$, where A is singular and $a = \operatorname{ind}(A)$ is known.

1. Pick x_0 and compute $r_0 = b - Ax_0$ and $A^a r_0$.
2. Compute $\beta = \|A^a r_0\|$ and set $v_1 = \beta^{-1}(A^a r_0)$.
3. Orthogonalize the Krylov vectors $A^a r_0, A^{a+1} r_0, \ldots,$ via the Arnoldi–Gram–Schmidt process carried out like the modified Gram–Schmidt process:
   For $i = 1, 2, \ldots,$ do
   Compute $h_{ji} = (v_j, Av_i),\ j = 1, \ldots, i.$
   Compute $\hat{v}_i = Av_i - \sum_{j=1}^{i} v_j h_{ji}$.
   Let $h_{i+1,i} = \|\hat{v}_i\|$ and set $v_{i+1} = \hat{v}_i / h_{i+1,i}$.
   (The vectors $v_1, v_2, \ldots,$ obtained this way form an orthonormal set.)
4. For $k = 1, 2, \ldots,$ form the matrices $\hat{V}_k \in \mathbb{C}^{N\times k}$ and $\bar{H}_k \in \mathbb{C}^{(k+1)\times k}$:

$$\hat{V}_k = [v_1|v_2|\cdots|v_k] \quad \text{and} \quad \bar{H}_k = \begin{bmatrix} h_{11} & h_{12} & \cdots & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & \cdots & h_{2k} \\ 0 & h_{32} & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & h_{kk} \\ 0 & \cdots & \cdots & 0 & h_{k+1,k} \end{bmatrix}.$$

5. Form the matrix $\hat{H}_m = \bar{H}_m \bar{H}_{m-1} \cdots \bar{H}_{m-a}$. (Thus, $\hat{H}_m \in \mathbb{C}^{(m+1)\times(m-a)}$ and is of the form given in (4.12).)
6. Compute the QR factorization of $\hat{H}_m$: $\hat{H}_m = Q_m R_m$; $Q_m \in \mathbb{C}^{(m+1)\times(m-a)}$ and $R_m \in \mathbb{C}^{(m-a)\times(m-a)}$. ($R_m$ is upper triangular.)
7. Solve the (upper triangular) system $R_m \xi_m = \beta(Q_m^* e_1)$, where $e_1 = [1, 0, \ldots, 0]^{\mathrm{T}}$.
8. Compute $x_m = x_0 + \hat{V}_{m-a}\xi_m$. (Then $\|A^a r_m\| = \beta\sqrt{1 - \|Q_m^* e_1\|^2}$.)

As can be seen from the above, the computation of the vector x_m requires that the $m - a + 1$ vectors $v_1, \ldots, v_{m-a}$, and x_0 (all of dimension N) be stored.

6. Application of DGMRES

Since the computation of the vector x_m requires $m - a + 1$ vectors of dimension N to be saved, it would be impractical to increase m indefinitely when applying DGMRES. (Exactly the same problem prevails also when applying GMRES to non-singular systems.) To remedy this problem, we propose to use DGMRES the way GMRES has been used in large-scale problems, namely, in the restarting (or cycling) mode. A description of this follows:

1. Pick the integer $m > 0$.
   Pick an initial approximation $u^{(0)}$ and set $x_0 = u^{(0)}$ and $i = 1$.
2. Starting with x_0, compute the vectors $x_1, \ldots, x_m$, by DGMRES.
   Set $u^{(i)} = x_m$.
3. If $u^{(i)}$ passes the accuracy test, stop and accept it as the solution.
   Otherwise, set $x_0 = u^{(i)}$, replace i by $i + 1$, and go to 2.

The steps that generate $u^{(i)}$ from $u^{(i-1)}$ form the ith cycle. This usage of DGMRES that generates the sequence $\{u^{(i)}\}_{i=0}^{\infty}$ will be denoted DGMRES(m), in analogy to GMRES(m). The advantage of this usage of DGMRES lies in the fact that we are able to achieve convergence using fixed moderate values of m, hence moderate storage.

It must be noted that if the vector $u^{(0)}$ in Step 1 of the algorithm above is such that $u^{(0)} = \hat{u} + \tilde{u}$ with $\hat{u} \in \hat{\mathscr{S}}$ and $\tilde{u} \in \tilde{\mathscr{S}}$, then, for all i, $u^{(i)} = \hat{u}^{(i)} + \tilde{u}^{(i)}$, with $\hat{u}^{(i)} \in \hat{\mathscr{S}}$ and $\tilde{u}^{(i)} = \tilde{u} \in \tilde{\mathscr{S}}$. (This follows from Theorem 3.1.) Thus, if we want to obtain only $A^{\mathrm{D}}b$, we can start with $u^{(0)} = 0$ in Step 1, which causes $\tilde{u} = 0$, and hence $u^{(i)} \in \hat{\mathscr{S}}$ for all i.

7. Numerical examples

Example 7.1. In our first example we take A to be a real $N \times N$ block-diagonal matrix, where $N = 2(n_1 + n_2 + n_3) + 5$, whose first $n_1 + n_2 + n_3$ blocks are of the form

$$\begin{pmatrix} a_k^{(i)} & b_k^{(i)} \\ -b_k^{(i)} & a_k^{(i)} \end{pmatrix}$$

while its last two blocks are

$$\begin{pmatrix} 0 & \epsilon_1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Therefore, A has $n_1 + n_2 + n_3$ pairs of complex conjugate eigenvalues of the form $a_k^{(i)} \pm \mathrm{i}\, b_k^{(i)}$. We distribute these eigenvalues in the complex plane as follows: The first n_1 pairs of them are located on the ellipse $\mathscr{E}_1 = \mathscr{E}(11, \mathrm{i}\sqrt{11}, 11)$, the next n_2 pairs of them are located on the ellipse $\mathscr{E}_2 = \mathscr{E}(11, \mathrm{i}\sqrt{11}, 3 + 2\sqrt{5})$ that is in the interior of $\mathscr{E}_1$, and the last n_3 pairs of the eigenvalues are situated on the degenerate ellipse $\mathscr{E}_3 = \mathscr{E}(11, \mathrm{i}\sqrt{11}, \sqrt{11})$ in the interior of $\mathscr{E}_2$. Here $\mathscr{E}(c, f, \rho)$ is the ellipse with center at c, foci at $c + f$, and sum of semi-major and semi-minor axes equal to ρ. Thus all three ellipses are confocal with center at $c = 11$ and foci at $c \pm f = 11 \pm \mathrm{i}\sqrt{11}$ and, therefore, have their semi-major axes perpendicular to the real axis. We choose $a_k^{(i)} = 11 + \alpha_i \cos(k-1)\theta_i$, and $b_k^{(i)} = \beta_i \sin(k-1)\theta_i$, $k = 1, \ldots, n_i$, where α_i and β_i are the semi-axes of the corresponding ellipse $\mathscr{E}_i$, and $\theta_i = \pi/(n_i - 1)$, $i = 1, 2, 3$. Next, we take the vector $\hat{b}$ to be the product $A\hat{x}$, where

Table 1
Application of DGCR (via DGMRES without cycling) to the singular system of Example 7.1[a]

m	$\|x_m - A^{\mathrm{D}}b\|$
3	6.32E + 00
5	4.59E + 00
7	3.22E + 00
9	2.09E + 00
11	1.24E + 00
13	6.85E − 01
15	3.46E − 01
17	1.53E − 01
19	6.06E − 02
21	1.85E − 02
23	5.16E − 03
25	1.46E − 03
27	2.46E − 04
29	1.79E − 05
31	1.27E − 06
33	1.85E − 08
35	5.51E − 10
37	4.72E − 10
39	4.45E − 10
41	4.32E − 10

[a]Here $x_0 = \cdots = x_3 = 0$, hence $\|x_3 - A^{\mathrm{D}}b\| = \sqrt{40} = 6.324\cdots$

$$\hat{x} = (\underbrace{1\ldots1}_{2(n_1+n_2+n_3)}\ \underbrace{0\ldots0}_{5})^{\mathrm{T}}.$$

Then we add to $\hat{b}$ a vector $\tilde{b}$ whose first $2(n_1 + n_2 + n_3)$ entries are zero and the last five are arbitrary. We take $x_0 = 0$ so that DGCR will produce purely $A^{\mathrm{D}}b$. Consequently, $x_0 = \cdots = x_a = 0$, and hence

$$\|x_a - A^{\mathrm{D}}b\| = \sqrt{2(n_1 + n_2 + n_3)}.$$

In Table 1 we present some of the results we obtained by applying DGCR (via DGMRES) to the system $Ax = b = \hat{b} + \tilde{b}$ with $n_1 = 10$, $n_2 = 5$, and $n_3 = 5$, and $\epsilon_1 = 1$ and $\epsilon_2 = 2$, hence $a = \mathrm{ind}(A) = 3$. The computations have been carried out in double-precision arithmetic. Note that the errors in x_m taper off at the end of the table but do not increase at least for several iterations. This can be viewed as evidence that DGMRES is a stable algorithm.

Example 7.2. Let M be an odd integer and denote by I and O, respectively, the $(M + 1)/2 \times (M + 1)/2$ identity and zero matrices. Define also the $(M + 1)/2 \times (M + 1)/2$ matrices T_1 and T_2 as in

$$T_1 = \begin{bmatrix} -2 & 0 & \dots & \dots & 0 \\ -1 & -1 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -1 & 0 \\ 0 & \dots & 0 & -1 & -1 \end{bmatrix},$$

$$T_2 = \begin{bmatrix} -1 & -1 & 0 & \dots & 0 \\ 0 & -1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & -1 \\ 0 & \dots & \dots & 0 & -2 \end{bmatrix}.$$

Finally, construct the $(M+1)^2 \times (M+1)^2$ non-symmetric matrix A as in

$$A = \left[\begin{array}{cccccccc|cccccccc} 4I & O & \dots & \dots & \dots & \dots & \dots & O & T_1 & -2I & O & \dots & \dots & \dots & \dots & O \\ O & 4I & \ddots & & & & & \vdots & -I & T_2 & -I & O & & & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & & & \vdots & O & -I & T_1 & -I & O & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & & \vdots & \vdots & O & -I & T_2 & -I & O & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & & \vdots & \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & \ddots & \vdots & \vdots & & & \ddots & \ddots & \ddots & \ddots & O \\ \vdots & & & & & \ddots & 4I & O & \vdots & & & & O & -I & T_1 & -I \\ O & \dots & \dots & \dots & \dots & \dots & O & 4I & O & \dots & \dots & \dots & \dots & O & -2I & T_2 \\ \hline T_2 & -2I & O & \dots & \dots & \dots & \dots & O & 4I & O & \dots & \dots & \dots & \dots & \dots & O \\ -I & T_1 & -I & O & & & & \vdots & O & 4I & \ddots & & & & & \vdots \\ O & -I & T_2 & -I & O & & & \vdots & \vdots & \ddots & \ddots & \ddots & & & & \vdots \\ \vdots & O & -I & T_1 & -I & O & & \vdots & \vdots & & \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots & \vdots & & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \ddots & O & \vdots & & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & O & -I & T_2 & -I & \vdots & & & & & \ddots & 4I & O \\ O & \dots & \dots & \dots & \dots & O & -2I & T_1 & O & \dots & \dots & \dots & \dots & \dots & O & 4I \end{array}\right].$$

The matrix A is obtained by first discretizing the Poisson equation

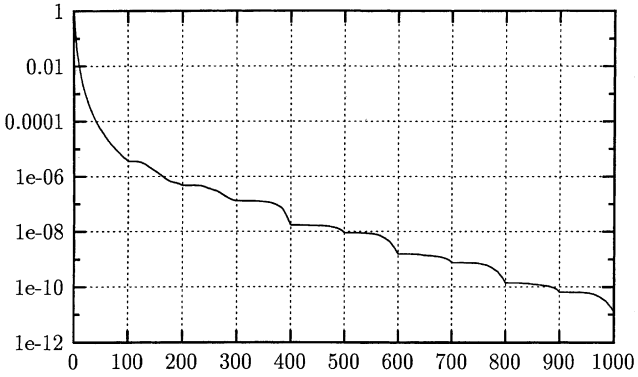

Fig. 1. Application of DGCR (via DGMRES(100) implementation) to the singular system of Example 7.2. Along the horizontal axis we give n, the number of matrix–vector products Av, while along the vertical axis we give $\log_{10}(\|z_n - A^{\mathrm{D}}b\|_\infty/\|A^{\mathrm{D}}b\|_\infty)$, where z_n stands for the DGCR approximation x_k in the $(i+1)$st cycle with $i = \lfloor n/100 \rfloor$ and $k = n - 100i$.

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) u(x, y) = f(x, y) \text{ for } (x, y) \in \Omega = [0, 1] \times [0, 1]$$

with Neumann boundary conditions

$$\frac{\partial}{\partial n} u(x, y) = \varphi(x, y) \quad \text{on } \partial\Omega\,,$$

on a uniform grid of mesh size $h = 1/M$ via central differences, and then by taking the unknowns in the red-black ordering. This problem was also considered by Hanke and Hochbruck [10] for testing the Chebyshev-type semi-iterative method that was developed there. Note that A is singular with a 1D null space spanned by the vector $e = [1, \ldots, 1]^{\mathrm{T}}$. Furthermore, $\mathrm{ind}(A) = 1$, as mentioned in [10]. Even if the continuous problem has a solution, the discretized problem need not be consistent. In the sequel we consider only the Drazin-inverse solution of $Ax = b$ for arbitrary b, not necessarily related to f and φ.

We first construct a consistent system with known solution $\hat{s} \in \mathscr{R}(A)$ via $\hat{s} = Ay$, where $y = [0, \ldots, 0, 1]^{\mathrm{T}}$. Then we perturb $A\hat{s}$, the right-hand side of $Ax = A\hat{s} = \hat{b}$, with a constant multiple of the null space vector e. In this way we end up with an inconsistent system with the Drazin-inverse solution $\hat{s}$. For this example our perturbation amounts to one percent in norm, i.e.,

$$\frac{\|\tilde{b}\|_2}{\|\hat{b}\|_2} = 0.01.$$

The initial vector x_0 is the zero vector.

In our numerical experiments we took $M = 63$. Therefore, the number of unknowns is 4096. With $M = 63$ the solution we are looking for is the vector $\hat{s}$, whose components are zeros except

$$\hat{s}_{2016} = -1, \quad \hat{s}_{2047} = -1, \quad \hat{s}_{2048} = -2, \quad \hat{s}_{4096} = 4.$$

Using DGMRES(100) in double-precision arithmetic, we obtain the results shown in Fig. 1. The results in this figure show quite clearly that errors in the x_m are decreasing constantly, which indicates that DGMRES in the cycling mode is useful.

Acknowledgement

The author would like to thank Vladimir Kluzner for programming DGMRES and for producing the graph shown in Fig. 1 of this work.

References

[1] W.E. Arnoldi, The principle of minimized iterations in the solution of the matrix eigenvalue problem, Quart. Appl. Math. 9 (1951) 17–29.
[2] A. Ben-Israel, T.N.E. Greville, Generalized Inverses: Theory and Applications, Wiley, New York, 1974.
[3] D. Calvetti, L. Reichel, Q. Zhang, Conjugate gradient algorithms for symmetric inconsistent linear systems, in: J.D. Brown, M.T. Chu, D.C. Ellison, R.J. Plemmons (Eds.), Proceedings of the Cornelius Lanczos International Centenary Conference, SIAM, 1993, pp. 267–272.
[4] S.L. Campbell, C.D. Meyer Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.
[5] J.J. Climent, M. Neumann, A. Sidi, A semi-iterative method for real spectrum singular linear systems with arbitrary index, J. Comp. Appl. Math. 87 (1987) 21–38.
[6] M. Eiermann, I. Marek, W. Niethammer, On the solution of singular linear systems of algebraic equations by semiiterative methods, Numer. Math. 53 (1988) 265–283.
[7] S.C. Eisenstat, H.C. Elman, M.H. Schultz, Variational iterative methods for non-symmetric systems of linear equations, SIAM J. Numer. Anal. 20 (1983) 345–357.
[8] R. Fletcher, Conjugate gradient methods for indefinite systems, in: A.G. Watson (Ed.), Numerical Analysis, Dundee 1975, Springer, New York, 1976, pp.73–89.
[9] R. Freund, N. Nachtigal, QMR: A quasi-minimal residual method for non-hermitian linear systems, Numer. Math. 60 (1991) 315–339.
[10] M. Hanke, M. Hochbruck, A Chebyshev-like semiiteration for inconsistent linear systems, Electr. Trans. Numer. Anal. 1 (1993) 89–103.
[11] C. Lanczos, Solution of systems of linear equations by minimized iterations, J. Res. N.B.S. 49 (1952) 33–53.
[12] Y. Saad, M.H. Schultz, A generalized minimal residual algorithm for solving non-symmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (1986) 856–869.
[13] A. Sidi, A unified approach to Krylov subspace methods for the Drazin-inverse solution of singular non-symmetric linear systems, Linear Algebra Appl. 298 (1999) 99–113.
[14] A. Sidi, V. Kluzner, A Bi-CG type iterative method for Drazin-inverse solution of singular inconsistent non-symmetric linear systems of arbitrary index, Electr. J. Linear Algebra 6 (1999) 72–94.
[15] H. van der Vorst, Bi-CGSTAB: A fast and smoothly converging variant of Bi-CG for solution of non-symmetric linear systems, SIAM J. Sci. Statist. Comput. 13 (1992) 631–644.